Partial Solution Set, Leon §5.3

5.3.1 Find least-squares solutions:

(a)

$$\begin{aligned} x_1 + x_2 &= 3 \\ 2x_1 - 3x_2 &= 2 \\ 0x_1 + 0x_2 &= 1 \end{aligned}$$

Solution: We are trying to solve $A\mathbf{x} = \mathbf{b}$, where $A = \begin{bmatrix} 1 & 1 \\ 2 & -3 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{b} = (3, 1, 2)^T$.

Clearly $\mathbf{b} \notin CS(A)$. So we use the normal equation, $A^T A\mathbf{x} = A^T\mathbf{b}$, which becomes

$$\begin{bmatrix} 5 & -5 \\ -5 & 10 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

The solution (unique, since A has rank 2) is $\hat{\mathbf{x}} = (2, 1)^T$.

(c)

$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \\ -x_1 + x_2 + x_3 &= 0 \\ -x_2 + x_3 &= 1 \\ x_1 + x_3 &= 2 \end{aligned}$$

Solution: The matrix equation is

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 4 \\ 0 \\ 1 \\ 2 \end{bmatrix},$$

which is inconsistent. The normal equations lead to the matrix equation,

$$\begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 6 \\ 3 \\ 7 \end{bmatrix},$$

so the solution is $\hat{\mathbf{x}} = (1.6, 0.6, 1.2)^T$.

5.3.2 For each solution $\hat{\mathbf{x}}$ in exercise 5.4.1,

    1. Determine $\mathbf{p} = A\hat{\mathbf{x}}$.

    2. Calculate $r(\hat{\mathbf{x}})$.

    3. Verify that $r(\hat{\mathbf{x}}) \in N(A^T)$.

For item (c) in 5.3.1, we have

$$\mathbf{p} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1.6 \\ 0.6 \\ 1.2 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 0.2 \\ 0.6 \\ 2.8 \end{bmatrix}, \text{ so } r(\hat{\mathbf{x}}) = \begin{bmatrix} 0.6 \\ -0.2 \\ 0.4 \\ -0.8 \end{bmatrix}.$$

We easily verify that

$$A^T r(\hat{\mathbf{x}}) = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.6 \\ -0.2 \\ 0.4 \\ -0.8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

5.3.3a Find all least squares solutions to $A\mathbf{x} = \mathbf{b}$, where $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ -1 & -2 \end{bmatrix}$ and $\mathbf{b} = (3, 2, 1)^T$.

Solution: First note that the columns of A are linearly dependent, so A (and hence $A^T A$) has a nontrivial nullspace and we anticipate multiple solutions. Solving

$$\begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 6 \\ 12 \end{bmatrix},$$

we find that all solutions have the form

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

5.3.4a For the system in Exercise 3, we want the projection $\mathbf{p}$ of $\mathbf{b}$ onto $R(A)$, and the verification that $\mathbf{b} - \mathbf{p}$ is orthogonal to each of the columns of A.

Solution: Continuing with the previous problem, the projection is

$$\mathbf{p} = A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

It follows that $\mathbf{b} - \mathbf{p} = (2, 0, 2)^T$, clearly orthogonal to the columns of A.

5.3.5a Find the best least-squares fit by a linear function to the given data:

$$\begin{array}{c|cccc} x & -1 & 0 & 1 & 2 \\ \hline y & 0 & 1 & 3 & 9 \end{array}.$$

Solution: We are assuming that $y = mx + b$, where m and b are the unknowns. Under this assumption, we have a system of equations,

$$\begin{aligned} 0 &= -m + b \\ 1 &= b \\ 3 &= m + b \\ 9 &= 2m + b, \end{aligned}$$

and the corresponding matrix equation is

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 9 \end{bmatrix}.$$

The normal equations become

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 21 \\ 13 \end{bmatrix}.$$

The solution: $m = 2.9, b = 1.8$, and so the function is $y = 2.9x + 1.8$.

5.3.6 Repeat problem 5.4.5a, but this time fit a *quadratic* polynomial to the data.

Solution: We now assume that $p(x) = ax^2 + bx + c$, where the coefficients are unknown. This leads to the equations,

$$\begin{aligned} p(-1) &= & a - b + c &= 0 \\ p(0) &= & c &= 1 \\ p(1) &= & a + b + c &= 3 \\ p(2) &= & 4a + 2b + c &= 9 \end{aligned}.$$

The corresponding (and inconsistent) matrix equation is

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 9 \end{bmatrix}.$$

The normal equations become

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix} = \begin{bmatrix} 13 \\ 21 \\ 39 \end{bmatrix}.$$

After elimination, we use back-substitution to find $a = 1.25$, $b = 1.65$, and $c = 0.55$, so $p(x) = 1.25x^2 + 1.65x + 0.55$.

5.3.9a Let A be an $m \times n$ matrix of rank n, and let $P = A(A^T A)^{-1} A^T$. Show that $P\mathbf{b} = \mathbf{b}$ for every $\mathbf{b} \in CS(A)$.

Proof: Let $\mathbf{b} \in CS(A)$. Then $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbf{R}^n$. It follows that

$$\begin{aligned} P\mathbf{b} &= A(A^T A)^{-1} A^T \mathbf{b} \\ &= A(A^T A)^{-1} A^T (A\mathbf{x}) \\ &= A(A^T A)^{-1}(A^T A)\mathbf{x} \\ &= A\mathbf{x} \\ &= \mathbf{b}, \end{aligned}$$

which is what we needed to show. $\square$

(Intuitively, the projection of $\mathbf{b} \in CS(A)$ onto $CS(A)$ must be $\mathbf{b}$ itself.)